

Qualidade de vídeo em uma arquitetura IPTV

Luciano Henrique Duque, engenheiro e mestre do UniCEUB - Centro Universitário de Brasília, na engenharia da computação

O artigo analisa a qualidade de vídeo (QoE - Quality of Experience) em função das características de perda de pacotes, taxa de erro e jitter, coletadas em uma arquitetura de rede IPTV com acesso ADSL2 e no limite de operação em relação à distância, conforme os valores recomendados pela norma WT-126. Com o resultado, é possível monitorar a infraestrutura IPTV de forma eficiente, gerando traps sempre que as métricas ultrapassarem os valores delineados.

O IPTV transmite sinais de multimídia, como áudio, vídeo, textos e gráficos, transportados em uma rede IP dedicada, com garantias de qualidade, confiabilidade e segurança. Nesse contexto, torna-se viável assegurar a qualidade de vídeo, garantindo uma característica no mínimo igual aos serviços existentes de TV.

O fator que determina a qualidade do vídeo é chamado de QoE - Quality of Experience. Aferir-lo em uma rede IPTV é uma tarefa complexa, porém, essencial para o controle e monitoramento da rede.

Converter perda de pacote, *jitter* e taxa de erro em medida de qualidade de vídeo é de suma importância para a rede IPTV, pois a coleta desses parâmetros é uma tarefa simples. O modelo de aferição de QoE numa rede IPTV é o rPSNR, associado a uma medida subjetiva de vídeo (MOS-V). Ele visa converter perda de pacote IP e *jitter* em qualidade de vídeo, partindo de um limite inicial definido pela recomendação da norma WT-126.

Arquitetura IPTV

A figura 1 mostra os elementos que compõe uma arquitetura de rede IPTV, baseada em tecnologia ADSL2.

Headend

Representa a extremidade principal de vídeo, ou seja, ponto no qual se encontra o conteúdo total de vídeo (filmes, programas, etc.), em conexões com operadoras de TV convencionais para a transmissão de programas ao vivo. O *headend* possui vários componentes que podem variar de rede para rede, incluindo fontes de vídeo analógico e digital, provedores de conteúdo, codificadores/decodificadores e transcodificadores para adaptar as taxas de *streaming*, switches, servidores para softwares de vídeo e de aplicação e servidores de gerência. No *headend* o vídeo será codificado em MPEG-4/H.264, com qualidade SD - Standard Definition a uma taxa de transmissão de 2 Mbit/s. O sinal de vídeo é processado com qualidade, sendo entregue ao backbone IP, onde todo o sinal é encapsulado por meio do protocolo IP e distribuído aos usuários.

Rede de núcleo IP

São redes preparadas para a transmissão de vídeo, garantindo a QoS - Quality of Service, que reflete na QoE aceitável pelo usuário. Sua qualidade é comparável a das TVs a cabo ou via satélite, podendo ser superior. Em

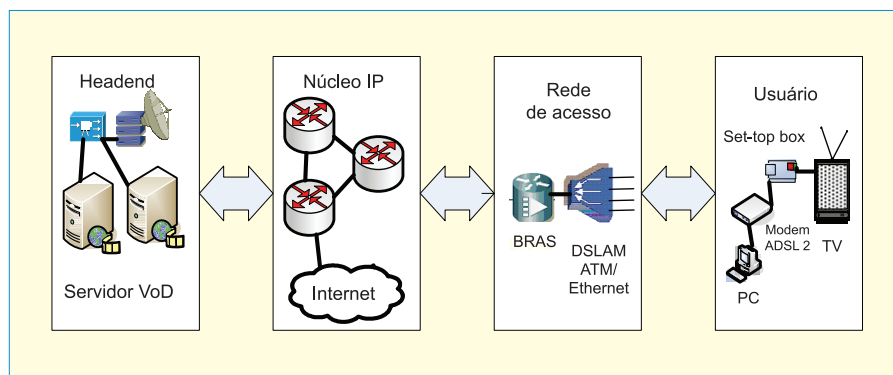


Fig. 1 – Componentes da arquitetura IPTV

termos gerais, o núcleo IP é uma rede, cuja estrutura física é baseada em fibra ótica ou redes de transporte, como DWDM. Ela agrupa os canais codificados de vídeo, transportando-os sobre a rede IP do provedor de serviço (backbone IP da operadora). O protocolo de transporte para o sinal de vídeo na arquitetura em análise é o UDP. O núcleo deve ser dotado de implementações de QoS, que possam garantir *jitter*, atraso e, principalmente, a perda de pacotes em limites aceitáveis, resultando em uma qualidade de vídeo satisfatória para o usuário final.

Rede de acesso

A rede de acesso faz parte da arquitetura de uma infraestrutura de IPTV, representando a conexão entre o fornecedor de serviço e a casa do usuário, ou seja, “a última milha”. A conexão do usuário pode ser realizada por meio de uma variedade de tecnologias de acesso - as operadoras vêm utilizando a tecnologia DSL. O DSLAM conecta os usuários através do par telefônico e sua saída pode ser ATM ou Gigabit Ethernet, dependendo da tecnologia empregada, em nossa análise utilizamos a Gigabit Ethernet. O DSLAM concentra os usuários e possui conectividade com o BRAS - Broadband Remote Access, que tem a função de fornecer o endereço IP ao usuário IPTV. Ele também tem o

endereço IP dos servidores de autenticação e redireciona sua autenticação a esses servidores, fazendo conexão com o núcleo IP,

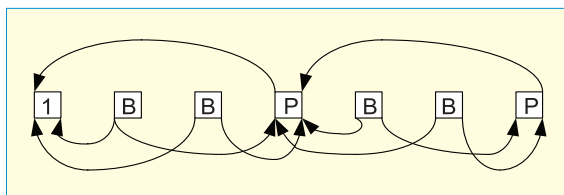


Fig. 2 – Formação dos quadros MPEG

para que o usuário busque seu conteúdo de vídeo no *headend*.

Ambiente do usuário

O modem ADSL2 permite velocidade de até 12 Mbit/s, em downstream, e 1 Mbit/s, em upstream, preservando o canal de voz.

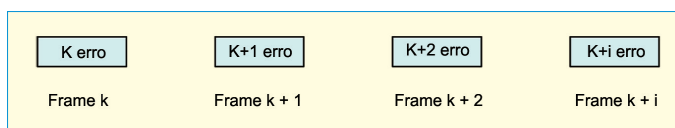


Fig. 3 – Sequência de vídeo

O set-top box é o elemento terminal de usuário que converte os *streams* para a saída composta de vídeo ou qualquer outra, de acordo com o padrão do aparelho de TV do usuário.

Valores recomendados pela WT-126 para a arquitetura em análise

O DSL Fórum recomenda valores de perda de pacotes, *jitter*, delay, etc., segundo a qualidade de vídeo que se deseja imprimir. A análise neste artigo será realizada para uma qualidade de vídeo em SDTV - Standard definition, utilizada pela Oi. A tabela I representa esses valores recomendados pela WT-126 e que serão utilizados, num primeiro momento, como valores-limite. Obviamente, iremos observar o comportamento da qualidade do vídeo e determinar os valores

máximos de perda de pacotes e *jitter*, onde o QoE ainda possa ser atendido, segundo as expectativas do usuário.

Modelo de RPSNR

Esse modelo define de forma relativa a qualidade do vídeo em função da perda de pacotes e taxa de erro. O sinal de vídeo compactado em MPEG-4/H.264 no *headend* é formado por quadros I, P e B. A figura 2 ilustra esses quadros.

Os quadros I não são dependentes dos quadros P e B, ao serem decodificados, mas são necessários para decodificação dos P e B. O quadro I representa uma imagem completa, codificada individualmente. Os quadros P são necessários para decodificação dos

quadros B e baseados em previsão antecipada. Os quadros P são codificados com previsão relativa ao último quadro P. O quadro B é a diferença entre o último e o próximo

Tab. I – Métricas de QoE, segundo a recomendação WT-126

Vídeo stream banda de transmissão (Mbit/s)	Latência (ms)	Jitter (ms)	Perda de período (pacote IP)	Perda de distância (evento errado a cada 30 min.)	Perda de pacote IP (E-6)
1,75	<200	<50	1	1	3,3
2	<200	<50	1	1	2,9
2,5	<200	<50	1	1	2,3
3	<200	<50	1	1	1,6

quadro P. Portanto, uma perda de pacote IP poderá provocar a perda de um frame I. Nesse caso, não será possível decodificar os quadros que chegam antes do próximo I.

Em relação às métricas de qualidade de vídeo, avaliaremos o MSE e o PSNR. O primeiro representa o erro quadrático médio entre uma imagem transmitida e uma imagem recebida.

O PSNR - Peak Signal to Noise Ratio representa a relação sinal/ruído MSE - Mean Square Error. Quanto mais próxima à imagem transmitida estiver a recebida, maior será o PSNR. A equação a seguir mostra o PSNR:

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} : \text{fonte [5]} \quad [1]$$

Ela representa o número de bits por amostra da imagem, esse valor geralmente é de 8 bits. O MSE é a diferença da imagem transmitida pela imagem recebida pelo usuário, conforme:

$$MSE = \left(\frac{1}{MN} \right) \sum_{i=1}^N \sum_{j=1}^M (u_{ij} - u'_{ij})^2 : \text{fonte [5]} \quad [2]$$

Aqui u_{ij} e u'_{ij} representam os pixels da imagem original e reconstruída, respectivamente.

Se ocorrer um erro em um frame k da sequência, indicado na figura 3, o erro tem o poder de propagar para o frame $(k+i)$, representado por $o2s$.

Logo podemos aproximar o erro quadrático médio da seguinte forma:

$$\sigma_s^2[k+i] = \sigma_s^2[k] \cdot \gamma^i : \text{fonte [5]} \quad [3]$$

Onde γ ($\gamma < 1$) representa o fator de atenuação do decodificador do set-top box e o i -ésimo frame. Assim, para um instante $(T-1)$ de codificação da imagem, os frames são codificados em P-frames, entre dois frames I-frames. A figura 4 ilustra a sequência de vídeo.

Vamos modelar o MSE, onde x simboliza o número de frames no ponto onde ocorreu a perda original até o próximo frame I, representando

uma variável aleatória, com distribuição uniforme de $[0, T-1]$. A figura 5 ilustra a variável x .

Podemos escrever o erro quadrático médio da seguinte maneira:

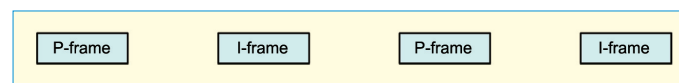


Fig. 4 – Codificação dos frames

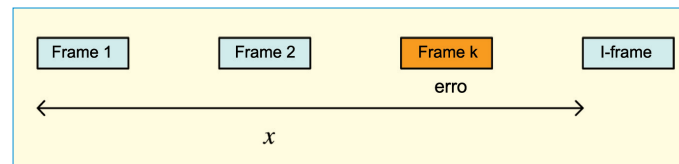


Fig. 5 – Representação da variável x

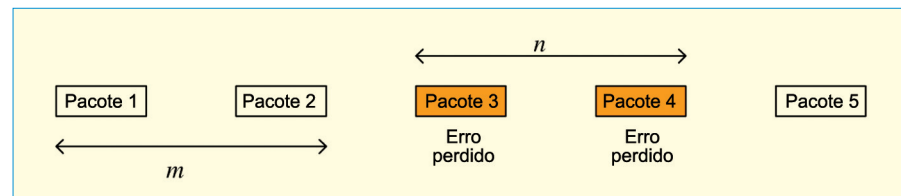


Fig. 6 Pacotes transmitidos entre dois eventos de perda



$$MSE = \sum_{i=0}^{x-1} \sigma^2[k+i]: \text{ fonte [5]} \quad [4]$$

Supondo que a distorção inicial é causada pela perda de uma fatia (pedaço) do frame, representado por uma constante (1 2s), logo o MSE é modelado da seguinte forma:

$$D_1 = \sum_{i=0}^{T-1} \sigma_s^2 \gamma^i \left(\frac{T-i}{T} \right) = \frac{\gamma^{T+1} - (T+1)\gamma + T}{T(1-\gamma)^2} \sigma_s^2 = \alpha \sigma_s^2 \quad [5]$$

Podemos afirmar que $\frac{1}{4}$ é função da atenuação e do tempo T. O valor de MSE representa o erro quadrático médio para a perda de uma única fatia do pacote. Como o erro propaga durante a transmissão, poderá existir mais de uma fatia perdida no pacote.

Quando tivermos uma perda de n pacotes consecutivos, sendo $n \leq 1$, teremos uma quantidade $f(n)$ de fatias perdidas do pacote, que faz o mapeamento de perdas de pacotes para perdas de fatia no quadro. Logo $f(n)$ irá depender do tipo de codificação, visto que estaremos trabalhando com a codificação MPEG-4. Vamos assumir s fatias por pacote e L pacotes por frame, considerando a perda uma distribuição uniforme de 1 pacote até L pacotes por frame. Assim, o MPEG-4 é:

$$f(n) = sn : \text{ fonte [5]} \quad [6]$$

Vamos modelar a distorção quadrática (MSE) para mais de uma fatia perdida por frame. Assim, o erro quadrático médio pode ser escrito como:

$$MSE = D_n = f(n)D_1 : \text{fonte [5]}$$

Estamos à procura de um modelo simples, onde medimos a perda de pacotes e convertemos em qualidade de vídeo. Definiremos m como

número de pacotes transmitidos entre dois eventos consecutivos com perda, ou seja, a distância onde ocorreu a perda (figura 6).

Denotamos P_n a probabilidade de termos n pacotes consecutivos perdidos. Já P_m representa a probabilidade de dois eventos consecutivos, separados por pacotes. Considerando m e n variáveis independentes, podemos determinar o MSE do IPTV reconstruído da seguinte forma:

$$MSE = \overline{D} = \frac{\sum_n P_n D_n}{\sum_m P_m (m/L)} = \frac{\overline{f(n)}}{\overline{m}} LD_1 : \text{fonte [5]} \quad [8]$$

A equação é equivalente a:

$$MSE = \overline{D} = P_e \overline{f(n)LD_1} : \text{fonte [5]}$$

[9]

Pe representa a probabilidade de ocorrer um evento de perda dentro de um stream de vídeo. Pe e $f(n)$

são em função das características do processo de perda visto no stream de vídeo, específico do codec utilizado. Por exemplo, se L é tipicamente grande, quando o vídeo é codificado em uma taxa elevada, ao substituir o valor de MSE da equação [9], teremos:

$$MSE = \overline{D} = P_e s n L D_1 : \text{fonte [5]}$$

Assim, computamos a qualidade de vídeo QoE em um trajeto da rede IPTV, usando a relação sinal ruído de pico:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) : \text{fonte [5]}$$

Porém, o valor estimado de PSNR requer não somente as estatísticas de perda na rede, mas também o conhecimento das características da aplicação. Contudo, o modelo que procuramos deverá depender apenas de coletas das características da rede IPTV. Para tanto, vamos estimar a métrica de um PSNR relativo, sem estimar o valor de DI . No primeiro momento devemos definir para a equação [9] a rede IPTV, as características mínimas de perdas de pacotes

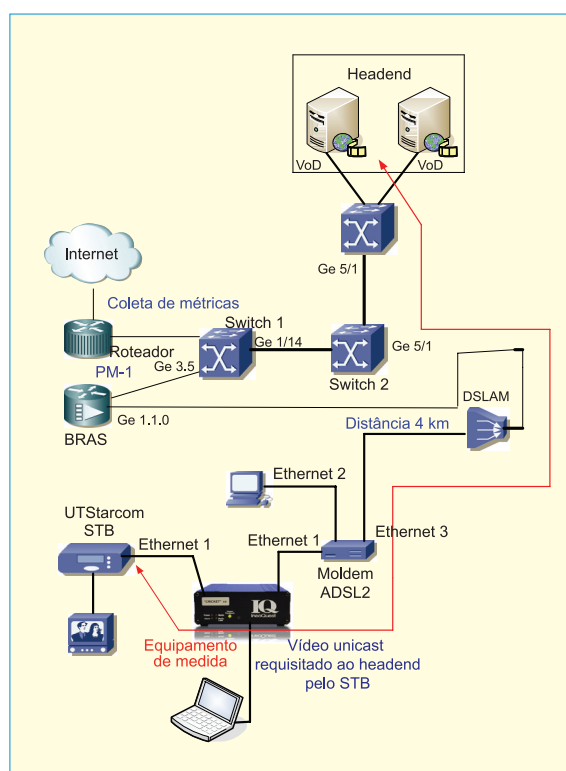


Fig. 7 – Arquitetura IPTV em medida

Tab. II – Tabela MOS-V

Nota MOS-V	Qualidade do MOS	Grau deficiência da imagem
5	Excelente	Imperceptível
4	Boa	Perceptível, mas não incômodo
3	Apropriado	Pouco incômodo
2	Pobre	Incômodo
1	Ruim	Muito incômodo

Tab. III – Valores recomendados pela WT-126

Vídeo stream banda de transmissão (Mbit/s)	Jitter (ms)	Perda de período	Perda de distância	Perda de pacote IP
2	< 50	1 pacote IP	1 evento errado a cada 30min	2,9E-6



e a probabilidade de erro para operação. Isto é, fazer $n = n_0$ e $P_e = P_{0e}$, baseados tipicamente em algum limite mais baixo da qualidade da rede IPTV, serem entregues ao usuário. Nesse caso, a qualidade relativa no trajeto IPTV é definida como a diferença do PSNR real e do PSNR alvo (PSNR transmitido dentro da rede IPTV). Logo, o rPSNR mede a distância que estamos da qualidade de vídeo esperada da rede. Escrevemos o rPSNR da seguinte forma:

$$rPSNR = 10 \log_{10} \frac{255^2}{D^0} - 10 \log_{10} \frac{255^2}{D} : \text{fonte [5]} \quad [12]$$

Aplicando a equação [9] na [12], sendo $D1$ o mesmo, obtemos o rPSNR final:

$$rPSNR = 10 \log_{10} \frac{(n^0)P_e^0}{(n)P_e} : \text{fonte [5]} \quad [13]$$

A equação [13] mostra que a qualidade de vídeo relativa pode ser estimada somente os valores de O , $n \cdot L$, P_{e0} e P_e , sendo n_0 e P_{e0} valores definidos pela WT-126. Esse é o modelo que será utilizado a QoE em uma rede IPTV. Nas simulações do trajeto da rede IPTV vamos utilizar o modelo de Gilbert [1], estimando a probabilidade de erro e perda de pacotes.

Medidas realizadas na arquitetura IPTV

As medidas foram realizadas na topologia mostrada na figura 7, nas seguintes condições:

- Vídeo compactado em MPEG-4/H.264.
- Tráfego de vídeo em CBR.
- Qualidade de definição para o vídeo em SD.
- Tráfego do vídeo em unicast.
- Taxa de compactação do vídeo 2,2 Mbi/s.
- Rede de acesso ADSL2.
- Distância do acesso ADSL2 até o headend de 4 km.

O equipamento é inserido entre o modem ADSL2 e o set top-box de forma serial. Ao se requisitar um

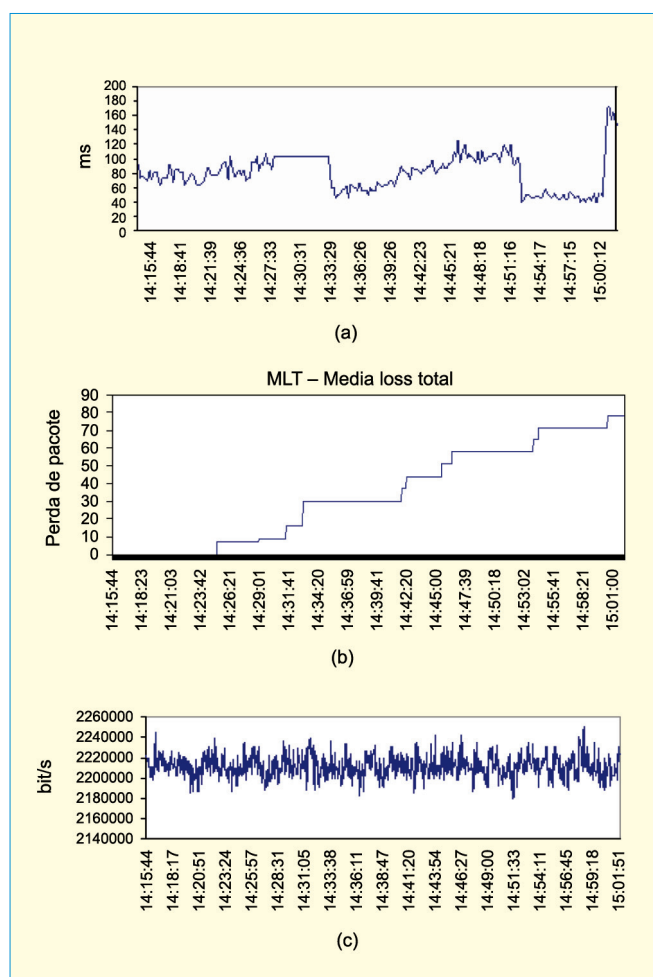


Fig. 8 – Medida de DF-delay factor (a), medida de MLR - perda de pacotes (b) e taxa de transmissão IP medida no acesso (c)

filme, o software captura o streaming de vídeo, analisando a perda de pacotes e seu *jitter*. A ferramenta ainda permite capturar pedaços de vídeo de até 90 Mb, que foi configurado para capturar sempre que ocorrer perda de pacotes IP. O vídeo medido teve uma duração aproximada de 45 minutos, período

do pacote de vídeo definido pela *iVMS* é possível determinar a taxa de erro no acesso ADSL2.

A perda de pacotes ocorrida no período de 30 minutos é a soma de todos os erros ocorridos, conforme a figura 9. O erro total é de 17 pacotes de média, assim a taxa de erro é calculada:

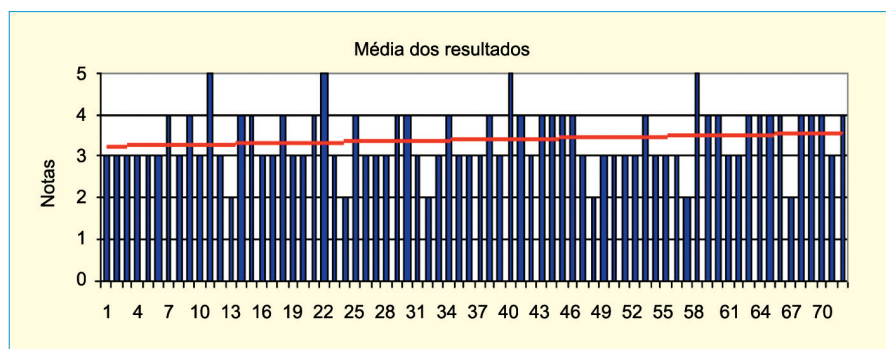


Fig. 9 – Notas do grupo entrevistado

em que a análise foi realizada. As medidas de *jitter* (DF- Delay Factor) e perda de pacotes (MLR - Midia Loss Rate) [2] são mostradas na figura 8.

Em 45 minutos, a perda média de pacotes foi de 3 e o *jitter* médio de 87,5 ms. O sinal de vídeo observado foi capturado pela ferramenta *iVMS* [3] e salvo em um arquivo formato *wincap*, para análise no Ethernet. Em função das taxas de perda de pacotes e de transmissão IP determinamos a taxa de erro.

Com o tamanho médio



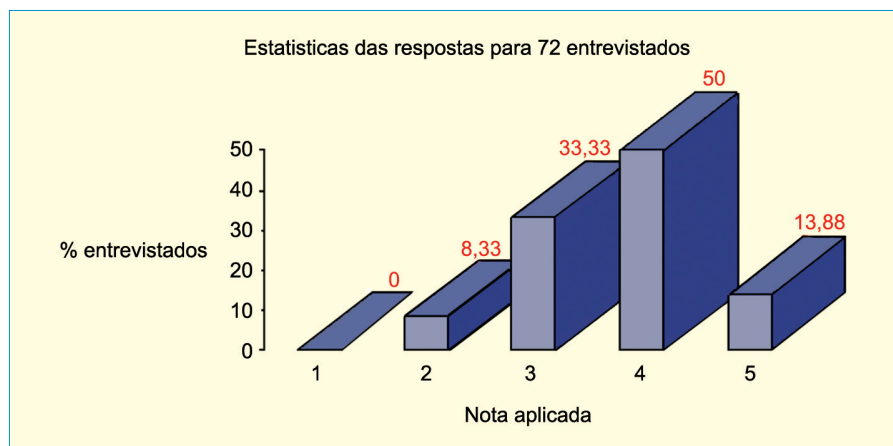


Fig. 10 – Nota aplicada em função da porcentagem de entrevistados

$$Tx_packet = \frac{Tx_rate \times time_collect}{packet_length} = \frac{(2,0 \times 1024 \times 1024 \times 30 \times 60) \div 8}{1310} = 360197.8625481 \text{ packets} \quad [13]$$

$$Rate_Error = \frac{Packet_error}{Packet_transmission} = \frac{17/7}{36019.78625481} = 6.74 \times 10^{-6} \quad [14]$$

Logo, podemos aplicar os valores medidos no modelo de rPSNR. Onde o valor esperado será o recomendado pela WT-126: $n=1$ e $Pe=2,9E-6$, conforme mostra a tabela I.

Observa-se que o rPSNR obtido foi negativo. Assim, o QoE não é atendido segundo o valor recomendado pela WT-126. O *jitter* médio medido foi de 87,5ms, conforme a figura 8. O vídeo analisado teve valores que não atenderam ao recomendado pela WT-126, quanto à perda de

pacotes, *jitter* e taxa de erro. Logo, o vídeo não atenderá aos requisitos de QoE para o usuário. Porém, o vídeo foi capturado e submetido a uma avaliação de um grupo de 72 pessoas, que deram notas, de acordo com a tabela II.

O grupo entrevistado tem o seguinte perfil

- Idade entre 18 e 40 anos.
- Estudantes de nível superior.
- Acesso à Internet.
- Conhecimento de serviços de TV por assinatura.

O resultado da pesquisa é apresentado na figura 9. As colunas representam as notas dadas pelos entrevistados e a linha vermelha a curva de tendência de notas para essa perda de pacotes. A nota média na observação desse vídeo foi de 3,4, representando uma qualidade entre apropriado e boa para a maioria dos usuários. A figura 10 mostra as estatísticas. Pela recomendação do WT-126, o

Tab. IV – Valores delineados para a arquitetura analisada

Vídeo stream banda de transmissão (Mbit/s)	Jitter (ms)	Perda de período	Perda de distância	Perda de pacote IP
2	< 87,5	2,43 pacotes IP	17 evento errado a cada 30min	6,74E-6

jitter, a taxa de erro e a perda de pacotes não atenderam aos requisitos de QoE. Porém, pela pesquisa MOS-V, o vídeo cumpriu os requisitos de QoE para a maioria dos usuários entrevistados.

Logo, foi delineado novos valores de perda, *jitter* e taxa de erro, como os limites máximos para a arquitetura de rede analisada, permitindo o atendimento mínimo aos requisitos de QoE. As tabelas III e IV, respectivamente, apresentam os valores recomendados pela WT-126 e os delineados para a arquitetura IPTV analisada.

Na realidade, é necessário atender uma qualidade máxima na rede IPTV, com erros nulos. Em se tratando de redes de acesso ADSL2, cujo canal é o par metálico, é provável que se conviva com erros. Portanto, delinear os erros máximos para garantir uma qualidade

mínima aceitável é de suma importância para o planejamento da cobertura do serviço IPTV.

Conclusões

Neste artigo analisamos a qualidade de vídeo em função das características de perda de pacotes, taxa de erro e *jitter*, coletadas em uma arquitetura de rede IPTV com acesso ADSL2 e no limite de operação em relação a distância. Os resultados de perda de pacotes foram aplicados no modelo rPSNR, traduzindo as perdas em qualidade de vídeo, associado ao *jitter* e a taxa de erro. Com os resultados é possível delinear as métricas máximas para a arquitetura analisada, possibilitando um planejamento da rede para a cobertura do serviço. Com isso será possível gerenciar a arquitetura de uma rede IPTV de forma eficiente,

gerando *traps* sempre que as métricas ultrapassarem os valores delineados.

Apesar de os valores atenderem ao grupo de usuários analisados, o objetivo, sempre que possível, é uma rede com perdas de pacotes e *jitter* estabelecidos pela WT-126. O artigo permite traçar em uma rede IPTV o perfil de qualidade e as áreas que podem ser atendidas pelo serviço.

REFERÊNCIAS

- [1] Documentação interna: *Sistema de Distribuição de Conteúdo*, SDC/IPTV, SEP-300724/2005-0100002135, Oi.
- [2] Request for comments 4445 (2006): *A Proposed Media Delivery Index (MDI)*. IneoQuest & Cisco Systems.
- [3] Software free Ethereal: <http://www.ethereal.com/download.html>.